# 6. MODEL EVALUATION

This chapter evaluates the model by analyzing and discussing potential sources of error, paying particular attention to the spatial data issues discussed in section 2.4.3; positional accuracy, resolution, and content. Studying the sources of model spatial error provides insight into developing validation studies and future research needs. One of the model objectives identified in chapter 3 was for the model to be statistically sound. Exploring the sources of error, and their propagation in the model, will also help determine appropriate strategies for developing confidence bounds around the spatially resolved estimates, an important model design feature. A sensitivity analysis and a comparison between the aggregate modal approach and speed-correction-factor approach are also provided.

A large amount of error in the model will be associated with the quality of the input data. While input data error is not explicitly discussed, it should be evident that any limitations associated with the input data impact the model results. It should also be noted that the input data's measures of spatial quality should focus on the relative positional accuracies among the datasets, not just the absolute accuracy.

## 6.1. Spatial Environment

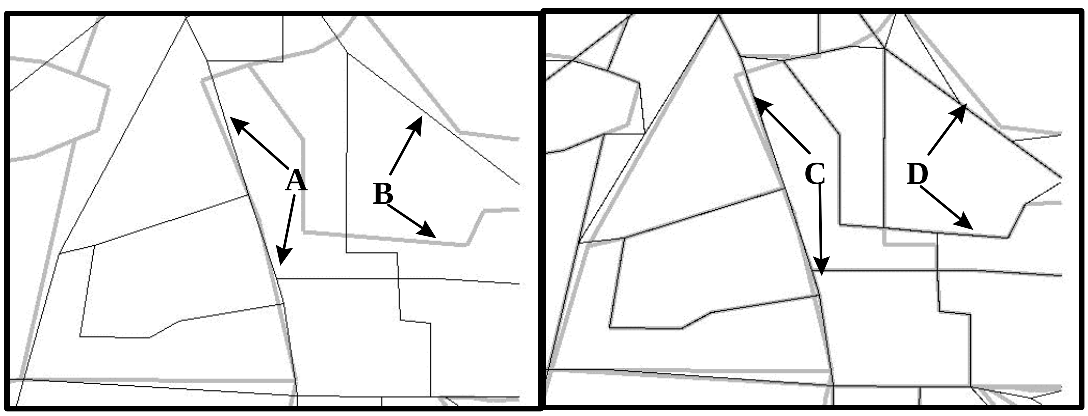
The spatial environment modules create the spatial entities *sz, mr,* and *mz.* Each entity was created by spatially manipulating input polygon and line data. During the spatial manipulation, potential positional errors arise that would impact the locational accuracy of the estimates. The following two sections describe the potential issues.

### 6.1.1. SZ

*SZ* was created using polygon-on-polygon overlay techniques on the four ARC/INFO coverages *census, taz, ZIPcode* and *landuse* (see figure 4.2). The technique merges two or more polygon networks into a single polygon network. The datasets share many common boundaries. However, the data were developed from different sources and resolutions resulting in different representations of common boundaries. In the sample area, the US Census blocks and the ARC's TAZs were generated from the original TIGER data, and, therefore, match very well. The ARC's land use and ZIP code data were developed from different sources. The impact of this problem is that there is potential for misrepresentation of the *landuse* / *TAZ* / *census* combinations.

Figure 6.1 demonstrates the polygon-on-polygon overlay problem. The figure shows a portion of the sample area's polygon structure. The left side shows *census* polygons in black lines and *landuse* polygons in gray lines. The right side shows *census* and *landuse* polygons in gray, and the resulting *SZ* polygons in black. Point A shows a shared boundary between *landuse* and *census* that is represented differently. Point B shows polygon boundaries that may, or may not, represent the same features; it is too difficult to tell conclusively.

The polygon overlay process includes a 'fuzzy tolerance' that allows the user to define a threshold that is allowable for matching edges. In the study model example, a tolerance of 30 meters was allowed. The resulting polygon structure (*SZ*) is shown on the right side of the figure. Point C shows the same area as point A, but the potential 'sliver' polygon was removed because the lines were within the tolerance level. Point D shows borders that may have represented the same feature, but as the distance between the two edges deviated more than 30 meters, they were represented as separate entities.




***Census* and *Landuse*** | ***Census*, *Landuse*, and *SZ***

**Figure 6.1 - Polygon Overlay Errors**

As a result of the polygon overlay process, there are potential errors that exist in the spatial representation of the joined polygons. There are three potential impacts on the resulting data; data are lost, data are spatially misrepresented, and data combinations are incorrect. The biggest risk to emission output quality is the loss of data. By adjusting the boundaries of the spatial entities, any polygons less than 30 meters across would be removed. In the sample study area used in Chapter 5, 5 of 925 US Census blocks were lost during this process. Most of the blocks were road medians between divided highways and held no bearing on the emission estimates.

However, one was a Census block that contained two households. In this case, two households have little impact, but other datasets may have more, suggesting that the model user would have to select a smaller tolerance level. The trade off to lower tolerance levels is increased spatial misrepresentation.

Spatial misrepresentation and incorrect data combinations are difficult to identify because the true values and positions must be known. Prior to model operation, detailed quality control steps in data development will prevent further error propagation. The model assumes that errors in input data exist, and therefore, a 'fuzzy tolerance' is used that can be adjusted to minimize data loss and maximize accurate representation.

### 6.1.2. MR and MZ

*MR* is created by selecting roads that are modeled in the travel demand forecasting model. Spatial errors associated with *MR* occur in the preprocessing steps, not in the formal model. The process of conflation, described in section 5.1.2, involves a great deal of user input, adding an aspect of human error.

The biggest concerns resulting from conflation errors are missing roads and miscoded roads. Matching some travel model links with the actual roads they represent can be difficult because the travel demand model network consists of abstract representations of roads. This is compounded by a lack of agreement or reporting about road classification in multiple datasets. Further, commercially available, accurate road datasets (similar to NAVTECH used in the study area) use significantly more detail than the travel models. One travel model link usually represents many road segments. During the conflation process, it is easy to miss one of the small segments, resulting in a 'gap' in the new network.

*MZ* polygons are created by defining the *MR* lines as polygon boundaries. There are no problems with the spatial accuracy of the polygons, other than those mentioned for the *MR* road network. However, there is one issue that is worth mentioning. The polygons are supposed to represent aggregations of local roads. The process of creating the *MZ* polygons does not actually consider the locations of these roads, but are defined as any polygon bounded by major roads (or the outside boundary). Therefore, medians from divided highways become defined as *MZ* polygons. While this poses no impact on the emission estimates, it can impact the amount of time required for model operation.

## 6.2. Vehicle Characteristics

The spatial errors associated with vehicle characteristics are significant and worth detailed study. There are several broad assumptions made in developing the spatially-resolved fleet distribution estimates. First, it is assumed that a vehicle's registered address is its 'home' location. This has not been proven or studied in previous research. Second, it is assumed that all the registered vehicles have the same probability of being operated at any given time. This not the case, but little evidence exists that justifies adjusting the fleet distribution to more accurately characterize operating vehicles. Third, it is assumed that any road segment's operating fleet distribution is composed of two groups of vehicles, a 'local' fleet and a 'regional' fleet. While some evidence exists to back up the idea [Tomeh, 1996], many questions remain about the specific definitions of the two groups of vehicles.

The potential negative impact of the above assumptions is reduced by predicting aggregate distributions rather than individual vehicles. The actual number of vehicles predicted at each entity, or whether the right vehicle is predicted at each entity, is less important than the predicted distribution. The only vehicle information used by the model is the fraction of each technology group at the zonal or road segment level, not the frequency. The model is more concerned with accurately characterizing the fleet, than accurately identifying the fleet.

Measures of the model's ability to predict the fleet distribution must come from future validation efforts. However, data do exist that indicate some biases found in the decoding process. Figures 6.2 and 6.3 show the degradation of the quality of the fleet estimate as vehicle characteristics are determined in the model. The basis for comparison (model year) comes from the raw vehicle registration dataset. As vehicle identification numbers (VINs) are decoded, model year information is predicted. Comparing the predicted and original model year distributions for the various datasets shows where bias has been introduced into the system.

Figure 6.2 shows the drop in the frequency of each model year for three steps in the decoding process; the VIN decoder, the removal of non-autos, and the assignment of vehicle weight. The VIN decoder operation results in a 7.7 % loss of data because the VIN couldn't be decoded. Most of those vehicles are older than 1980, when VINs were not standardized among manufacturers. Two odd 'humps' occur (frequency is overpredicted) in model years 1973 and 1978. After removing non-autos, only the 1973 hump remains. Further study revealed that the VIN decoder software was incorrectly assigning pre-1972 BMWs and Volvos as 1973 vehicles.

Adding the test weight to the vehicles (removing those without matches) resulted in a substantial data loss (39%). It also appears that the data loss is biased by model year, with 1988 vehicles underrepresented and 1995 and newer vehicle unrepresented. Figure 6.3 shows the resulting distributions. The final distribution shows the impact of the data loss on the fleet distribution. Pre-1972, post-1994, and 1988 vehicles are under-represented. Mid-1980s and early-1990s vehicles are over-represented.

### 6.2.1. Zonal Fleet

There are two concerns regarding the spatial allocation of the vehicles to zones; incorrect assignment, and non-residential trip distributions. Vehicles can be incorrectly assigned to zones because of address-matching problems. Non-residential trips use the fleet distribution of the current zone, disregarding the fact that the actual fleet distribution consists of vehicles originating from other locations.

The address-matching process can result in a small percentage of vehicles that were incorrectly assigned to zones because of errors in the address, an issue that has been well-documented in the literature. This problem is minimized in the model by having stringent matching guidelines; the vehicles' ZIP codes must match the candidate addresses' ZIP codes, the road types must match perfectly, and there can only be one error in spelling or incorrect address prefix (north, east, etc.). Further, the road dataset being used for address-matching could be missing new subdivisions or developments. If the registration dataset is newer than the last road dataset update, there could be vehicles that fail to match. However, the 'failed' vehicles are not discarded, but assigned a location based on their ZIP code.

The zonal fleet is developed from two sets of files, an address-matched file, and a ZIP code file (address match failures). To bring these two groups of vehicles together, the relationship between the *SZ* polygons and *ZIPcode* polygons must be identified. Each *SZ* is apportioned part of the ZIP code vehicles based on a comparison of the areas of the two polygons. A zone could have 78 vehicle address-matched within its boundaries, and an additional 10.3 vehicles assigned to it from the ZIP code. Since the concern of the model is the distribution, there are no problems that arise from non-integer frequencies.

Zones that do not have any address-matched vehicles are assigned the fleet distribution of the ZIP code. While some problems remain, zones that have new subdivisions will be assigned a fleet distribution that partially represents the vehicles registered at that location.

The issue regarding the fleet distribution of non-residential trips is not handled well in the model. Since vehicles are not tracked during estimates of activity, there is no mechanism for tying the origin fleet distribution to the destination. Unless the destination lies in a zone or ZIP code with a fleet distribution that is similar that of the origins, an incorrect distribution will be assigned. Given the dynamics of land use development, there is strong indication that strong bias will exist. For example, the fleet distribution of vehicles leaving a commercial land use zone is assigned the fleet profile of registered vehicles in that zone and ZIP code, not the trip origin zones of the operating vehicles.

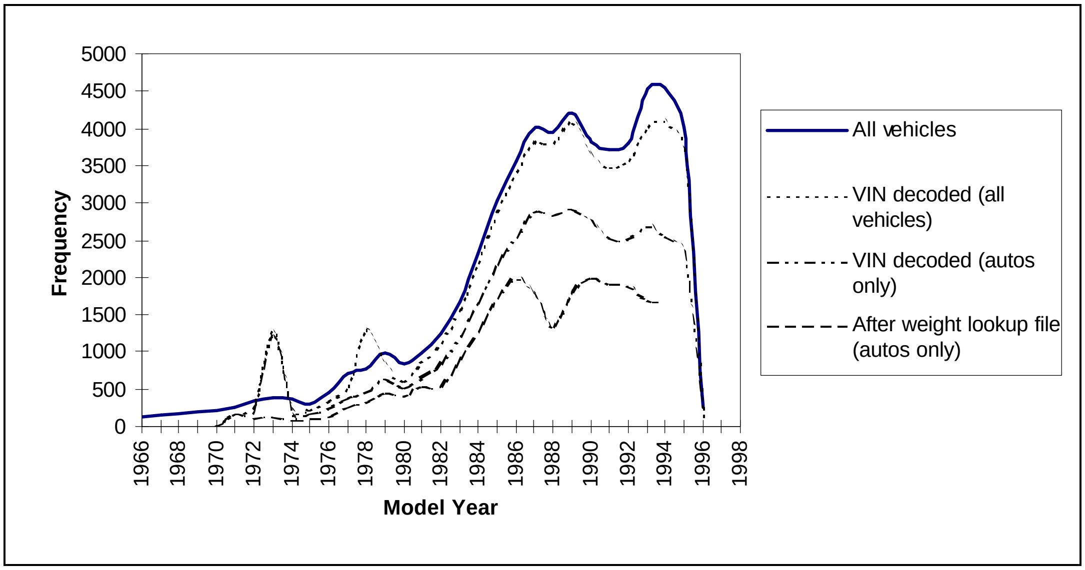



**Figure 6.2 - Model Year Frequencies**

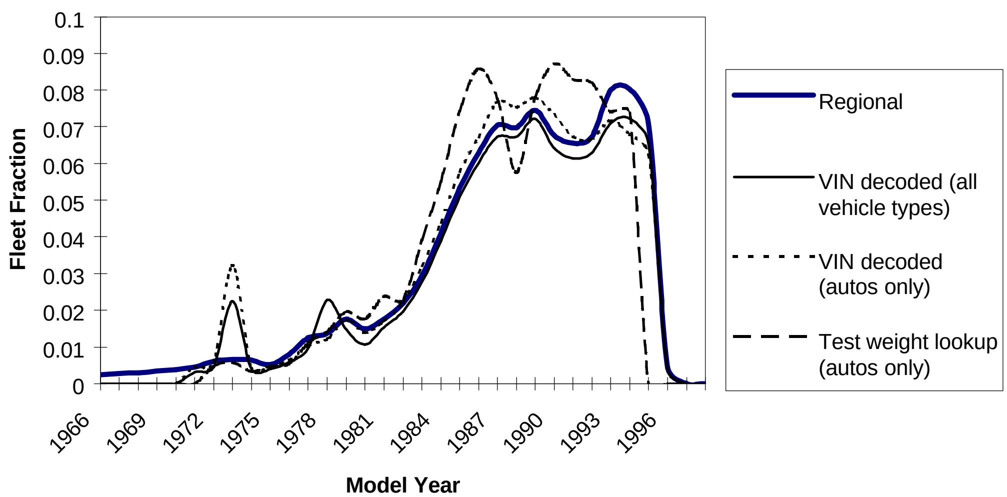
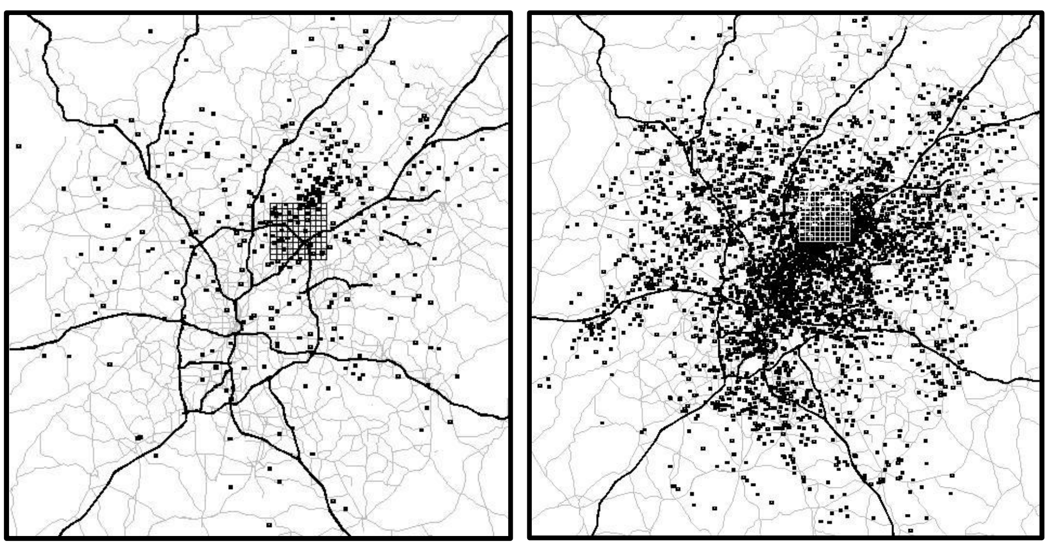



**Figure 6.3 - Model Year Fraction**

### 6.2.2. On-road Fleet

Problems associated with the on-road fleet distribution estimation stem from the unvalidated assumption that the on-road fleet can be summarized by combining a local fleet (defined as all vehicles within 3 km) and a regional fleet (all vehicles in the region). The instability of this approach is demonstrated by analyzing two sets of observed on-road vehicle datasets. Figure 6.4 shows registered vehicle locations for vehicles that passed through data collection sites in the study area. These data were collected and provided by the School of Earth and Atmospheric Sciences at Georgia Tech. License tags were captured on passing vehicles and matched to a registered vehicle's VIN and address. At site A, 674 vehicles passed through the data collection site. The figure indicates that spatial variation exists in the estimated origins of the observed vehicles. However, the size and shape of the spatial variability are unclear. Similarly, site B, with 13,481 vehicles, indicates spatial variability. The model currently uses a 3 km radius to define a 'local' fleet (10% of the observed vehicles fell in that range). It may be more accurate to select an alternative geometric (wedge, oval, network distance, etc.) search pattern involving road types, time of day, and network structure.



Site A: Arterial, 641 Vehicles

Site B : Ramp, 13481 Vehicles

**Figure 6.4 - Observed On-Road Vehicle Origins**

## 6.3. Vehicle Activity

Spatial errors associated with the estimates of vehicle activity can be tied to previously mentioned problems with the spatial environment and the travel demand model limitations. The model shares the problems associated with the use of the travel demand forecasting models in predicting emission-specific vehicle activity; inaccurate speeds, no feedback into the distribution phase, etc. (see section 2.3.1). These known travel demand model result limitations will not be discussed. However, trip disaggregation, the use of regional temporal distributions, and speed and acceleration matrices create some errors in vehicle activity estimates that are worth mentioning.

The disaggregation of trips by purpose to different landuse makes the broad assumption that the landuse data are discrete. All home-based trips are assigned origin engine starts to residential area. If homes can be found in other land uses, their engine start activity estimates will not be assigned to the correct location. The landuse data must be discrete to prevent this from occurring.

The use of *regional* temporal factors to distribute *zonal* and *road segment* activity results in errors. A series of spatial queries using 1990 Census data in Atlanta indicated that the fraction of people traveling to work between 6:30 and 7:00 AM was approximately 9% for people living within 2 kilometers of the central business district (CBD), and about 15% for people living between 8 and 10 kilometers from the CBD. By using regional temporal factors in the model, all zones and road segments are assigned the average, not allowing spatial variability. Thus, the peak hour 30 miles from the CBD will be the same as the peak hour 5 miles from the CBD.

The speed and acceleration profiles were used as a post-processor to the travel demand model's output. They are used to predict the modal distributions of the vehicles operating on the road. The model only includes matrices for interstates and ramps, forcing all lower classifications to rely on a single profile of mid-block estimates. As soon as new data are collected, validated, and available, the model structure can incorporate the new findings. The impacts of modal activity around signalized intersections could have a tremendous impact on the spatial variability of the estimates. As is, running exhaust emission estimates are highly correlated to volume. The potential variability found in future matrices could show that the highest emissions occur around major intersections, not high volume, low modal variability interstates. Further, the characterizing of dynamic modal activity into discrete bins and levels of speeds and accelerations could result in a certain level of error. Current research efforts are attempting to validate the approach.

## 6.4. Facility and Gridded Emissions

The spatial errors associated with the emission estimates come from aggregation. No spatial manipulation procedures are used to generate the facility-level emissions estimates. The facility-level estimates are, however, impacted by non-spatial errors. The gridded emission estimates are generated by aggregating facility-estimates to vector grid cells of a user-defined size. During this process, spatial errors are incurred.

### 6.4.1. Facility Emission Estimates

New errors introduced to the facility-level emission estimates are generated by the process for determining emission rates. The emission modes (engine start, running exhaust) have gram per start or gram per second rates predicted by the hierarchical tree-based regression. The resulting emission rate values are discrete, with known confidence bounds. The accuracy of the emission rates is affected by the size and representativeness of the emission test dataset. The emission rates used in the model were developed from a dataset of approximately 3000 vehicle tests using about 700 individual vehicles. Currently, improvements and additions are expanding the dataset to over 10,000 tests.

### 6.4.2. Gridded Emissions

The errors associated with the aggregation of facility emissions to user-defined grid cells are spatial in nature. Vector grid cell polygons are overlaid with the *SZ, MR*, and *MZ* entities using the techniques described in sections 6.1.1 and 6.1.2. The 'fuzzy tolerance' used in this process is set as low as processing time will allow. There won't be any shared boundaries in this overlay technique and high 'fuzzy tolerance' will only degrade the spatial quality. Once the polygons and lines are split by the grid cells, the emission estimates are weighted (the proportion of the new entities' area or length and the original spatial entities' area or length) and summed by grid cell.

The size of grid cell selected by the user impacts the cell's accuracy. Larger cells will have more accurate estimates because errors (unbiased error) at the facility levels can be offset by aggregation. Small grid cells will have fewer entities falling within its borders, reducing the number of values to draw from. Further, grid cell sizes falling below the spatial accuracy of the origin datasets could spatially misrepresent the locations of emission estimates. Larger cells have the advantage of absorbing errors related to absolute position. Figure 6.5 shows grid cell aggregations from the sample study area described in chapter 5. Four levels are shown: 100 meter, 250 meter, 500 meter, and 1000 meter. The figure is useful in looking at the total emissions from different levels of aggregation. While the 1000 meter (1 km) grid cell is expected to be used for future photochemical models, additional information for research can be gleaned from smaller cell sizes.

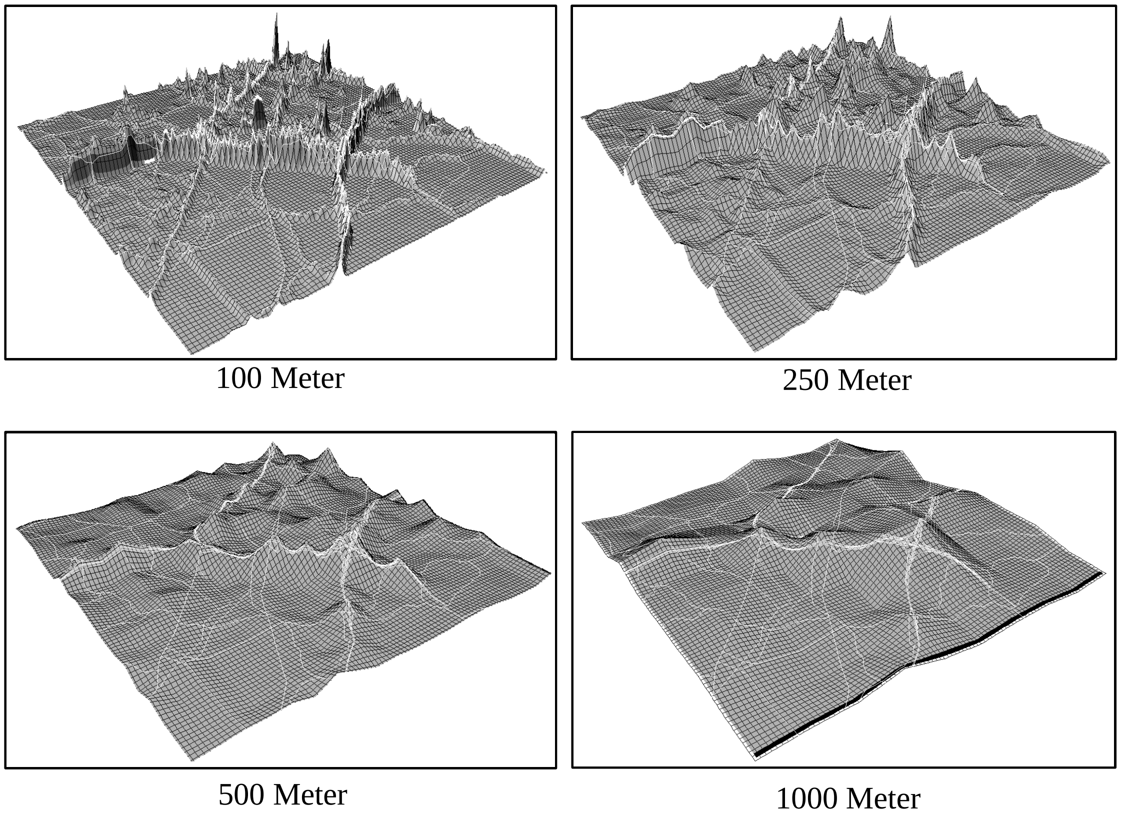



**Figure 6.5 - Sample Grid Cell Aggregations**

### 6.4.3. Sensitivity of Model

The model sensitivity can be measured in two ways: estimate accuracy and locational accuracy. The sensitivity of the estimate accuracy can be shown by running the model with a full range of input variables. The sensitivity of the locational accuracy depends on the spatial allocation of the estimate, given a full range of influential factors.

Figures 6.6 through 6.11 show how the emission rate varies for each technology group, level of service (LOS), and road grade. The graphs are for interstate activity only because speed and acceleration data for lower classifications do not yet exist. The percentage of the sample area's regional fleet in each technology group is provided in the graph as well. The very low percentage of some technology groups is the result of the problems mentioned earlier (section 6.2) regarding the determination of vehicle technologies.

All the technology groups have substantial estimated increases in emission rates for LOS F. The speed and acceleration profiles for interstate LOS F show substantially more variability in speeds and accelerations. Other LOS impacts are fairly static, slightly increasing as traffic flow degrades from LOS A to LOS D. However, during LOS F activity (volume to capacities greater than 1.0 and average speeds less than 30 MPH), the model is predicting that emissions rates increase substantially.

The impact of road grade is seen for CO normal emitters, HC high emitters, and NOx high emitters. The graphs indicate that these technology groups have higher emission rates for steeper grades. As mentioned previously, the impacts of grade may be substantially under-predicted. Currently, the model adjusts the acceleration rates of vehicles based on the road grade. There are no mechanisms in the model that adjust emission rates based on engine load. It is expected that emission rates will vary significantly once these impacts are considered.

Unlike the emission estimates, the locational sensitivity of the model is not the result of a series of calculations. Estimates are allocated to zones or lines based on input data conditions. For example, the return trip of a home-based-shopping (HBSH) trip begins in a shopping area (commercial land use) and ends at home (residential land use). If the TAZ with a HBSH attraction has commercial land use within its boundaries, the emissions from the engine start are allocated evenly to all commercial areas. If no commercial land use is indicated by the data, the engine start emissions are all allocated evenly to the entire TAZ. All the sample area TAZs had residential and non-residential land uses, and all but two had commercial land uses.

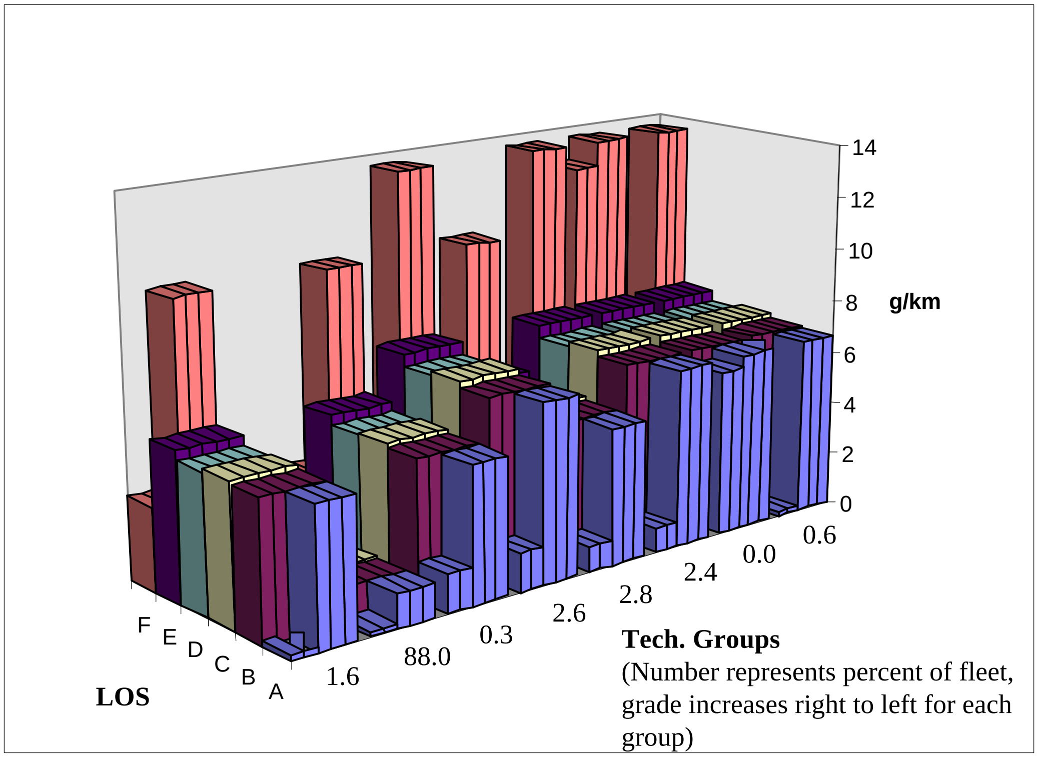



**Figure 6.6 - CO normal emitter technology group emission rates by LOS and grade**

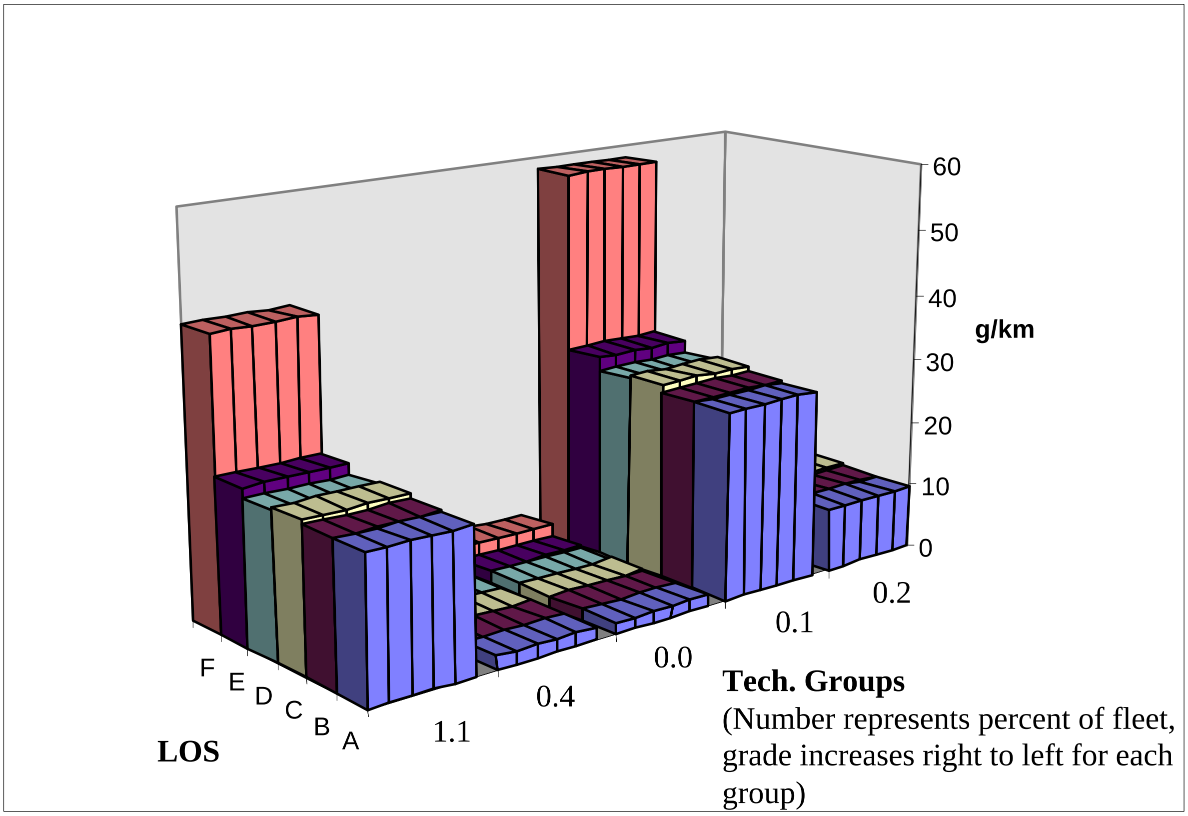



**Figure 6.7 - CO high emitter technology group emission rates by LOS and grade**

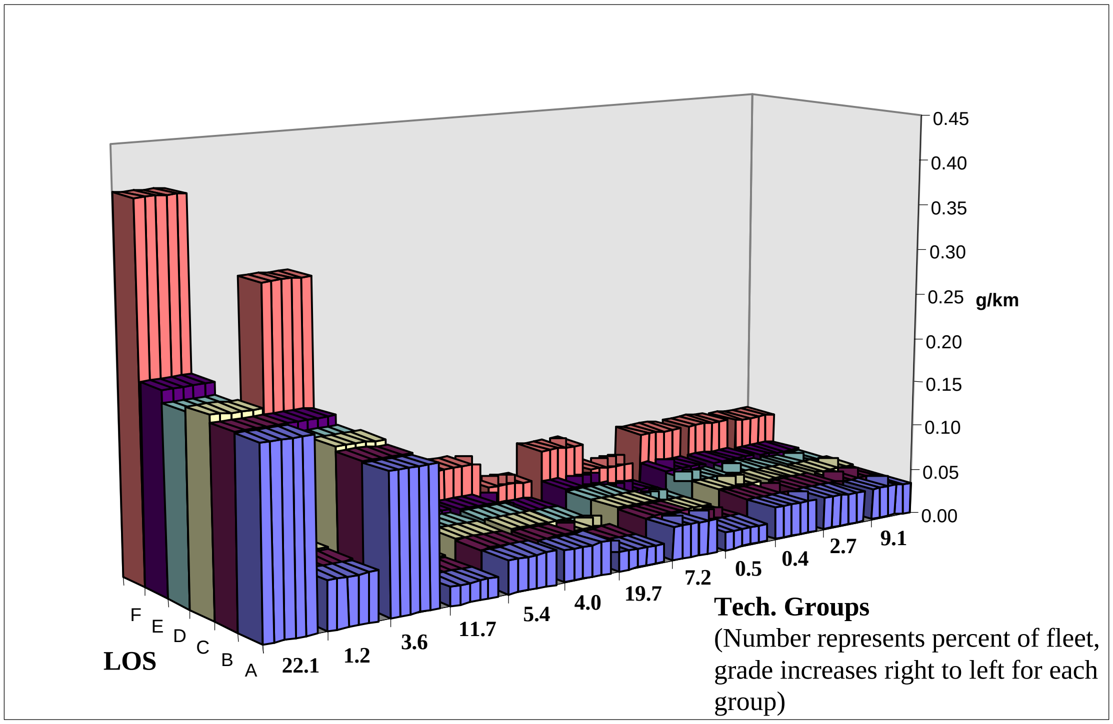



**Figure 6.8 - HC normal emitter technology group emission rates by LOS and grade**

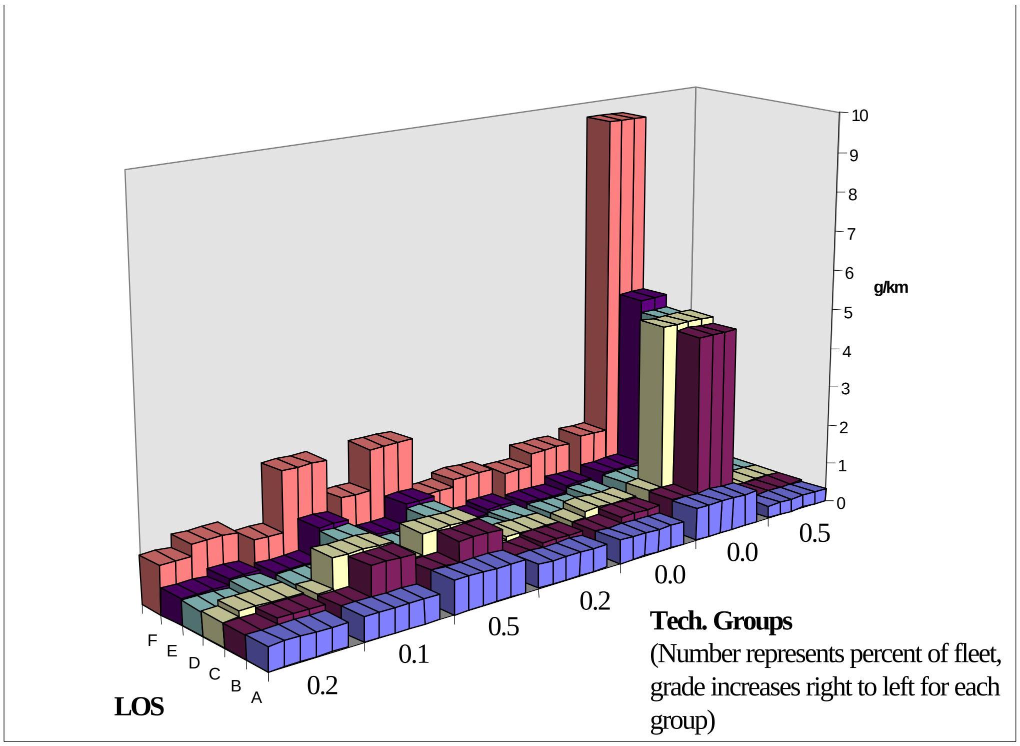



**Figure 6.9 - HC high emitter technology group emission rates by LOS and grade**

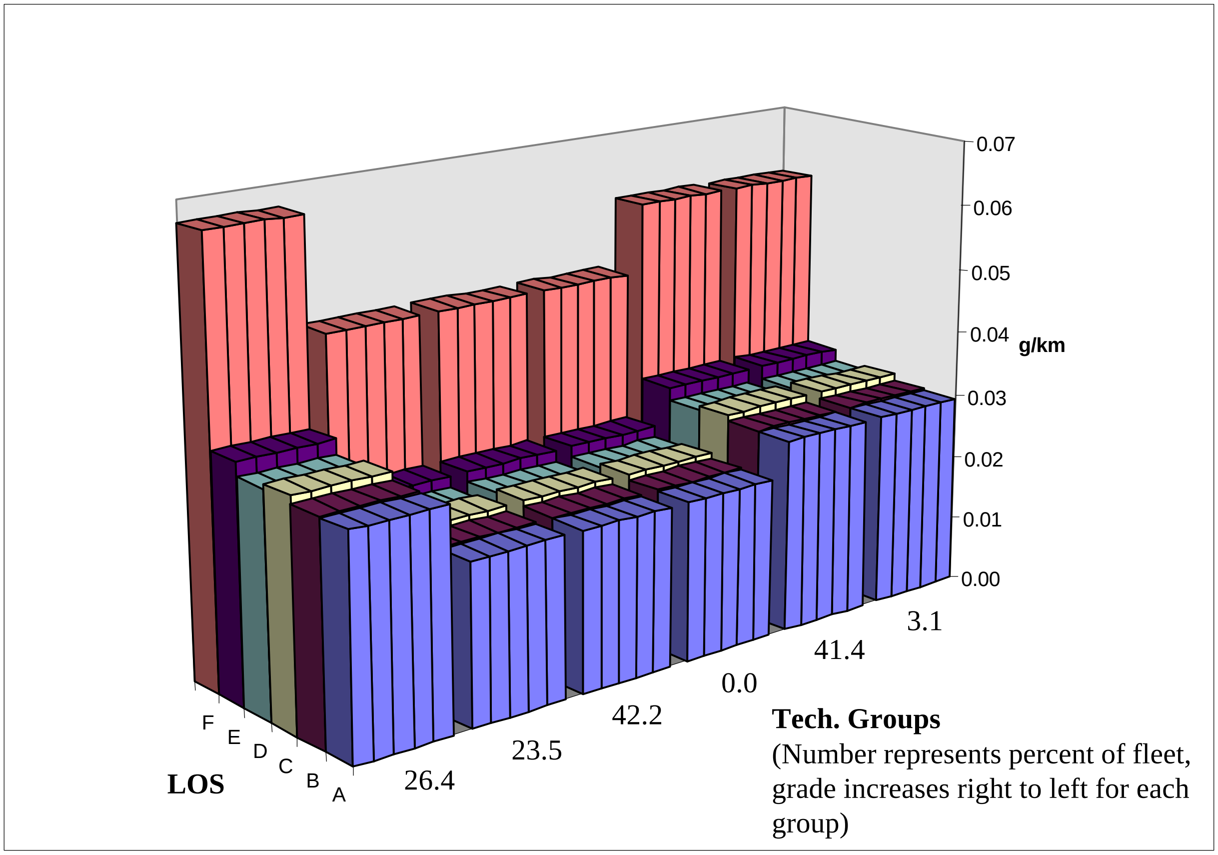



**Figure 6.10 - $NO_x$ normal emitter technology group emission rates by LOS and grade**

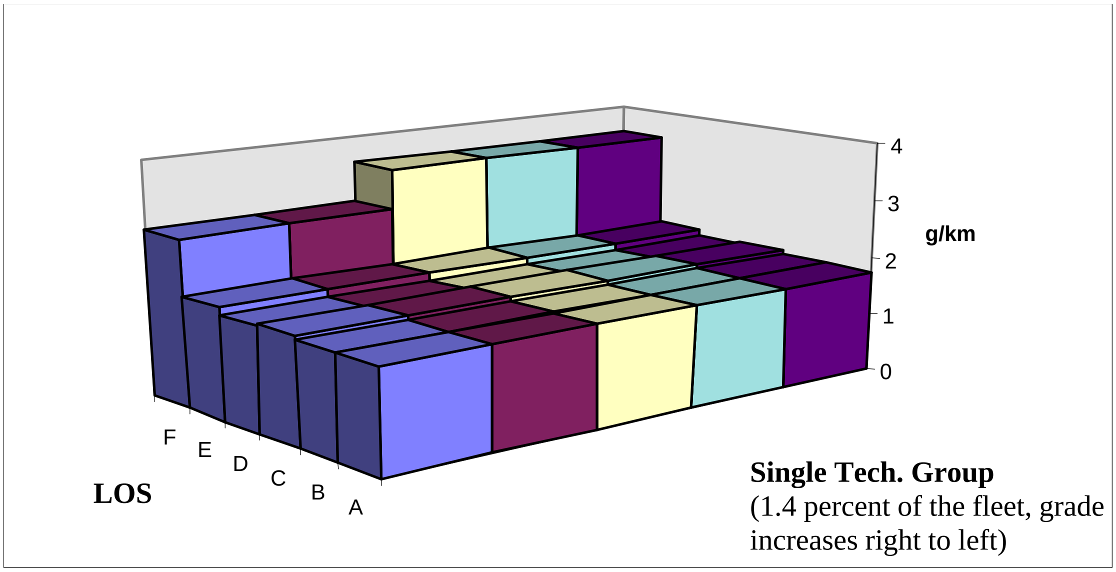



**Figure 6.11 - NOx high emitter technology group emission rates by LOS and grade**

### 6.4.4. MEASURE vs. MOBILE5a

To compare the USEPA's MOBILE5a and the new HTBR emission rates used in MEASURE, emission rates were determined for each speed and acceleration bin (0-80 mph, -10.0 to 10.0 mph/sec). Figures 6.12 to 6.17 show these profiles. Both emission rate models were used for each pollutant. The sample area results are also provided, showing a comparison between the hourly total grams of each pollutant. While much remains to be validated with MEASURE, this comparison provides some evidence for the future development of modal emission rate models.

Both data sets used in the analysis included the regional fleet distribution for the study area. The MOBILE5a rates were the running exhaust zero mile base emission rates (deterioration and start fractions effects removed). The HTBR rates used in MEASURE were similar; no start emissions or deterioration effects were included.

All the graphs show significant differences. The biggest impacts is the fact the MOBILE5a does not vary emissions by acceleration. A vehicle traveling at an average speed of 50 mph with minor variations in acceleration and deceleration is predicted to have the same emission rate as one with large variations. MEASURE indicates that these variations may have significant impacts on emission rates at certain thresholds of speed and acceleration activity.

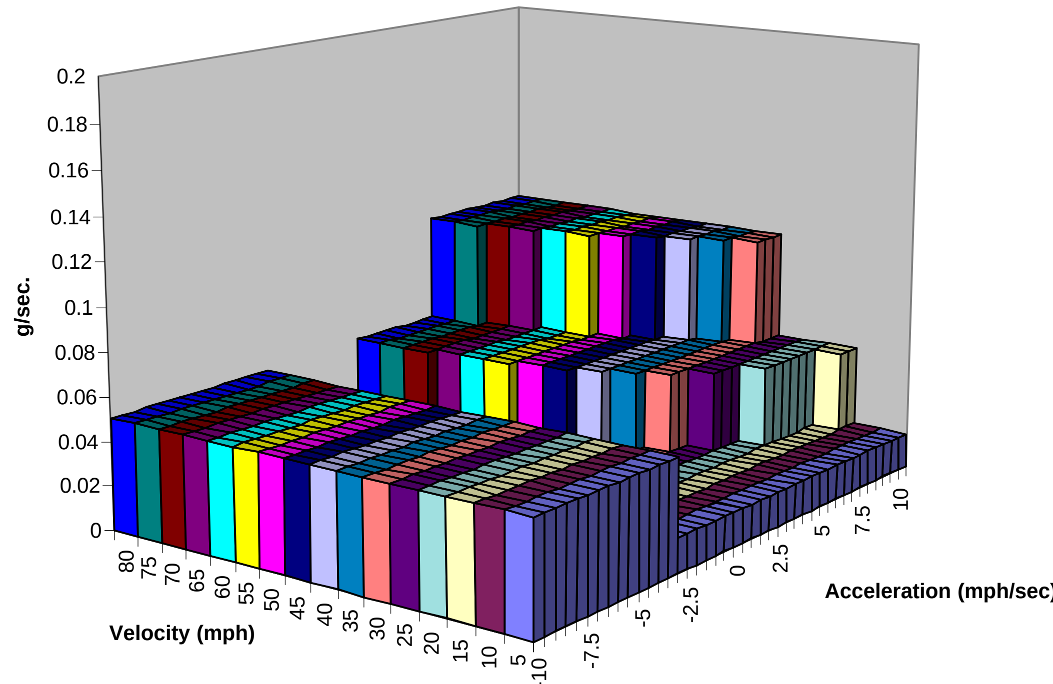



**Figure 6.12 - MEASURE g/sec CO emission rates by velocity and acceleration for the study area's vehicle fleet**

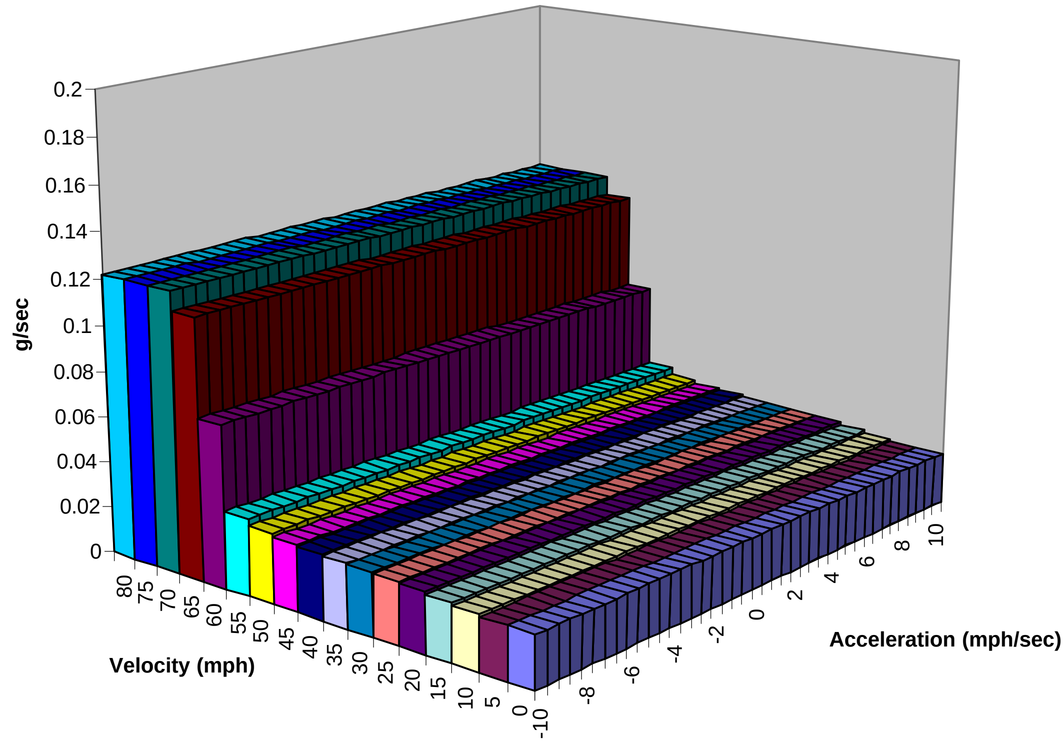



**Figure 6.13 - MOBILE5a g/sec CO emission rates by velocity and acceleration for the study area's vehicle fleet**

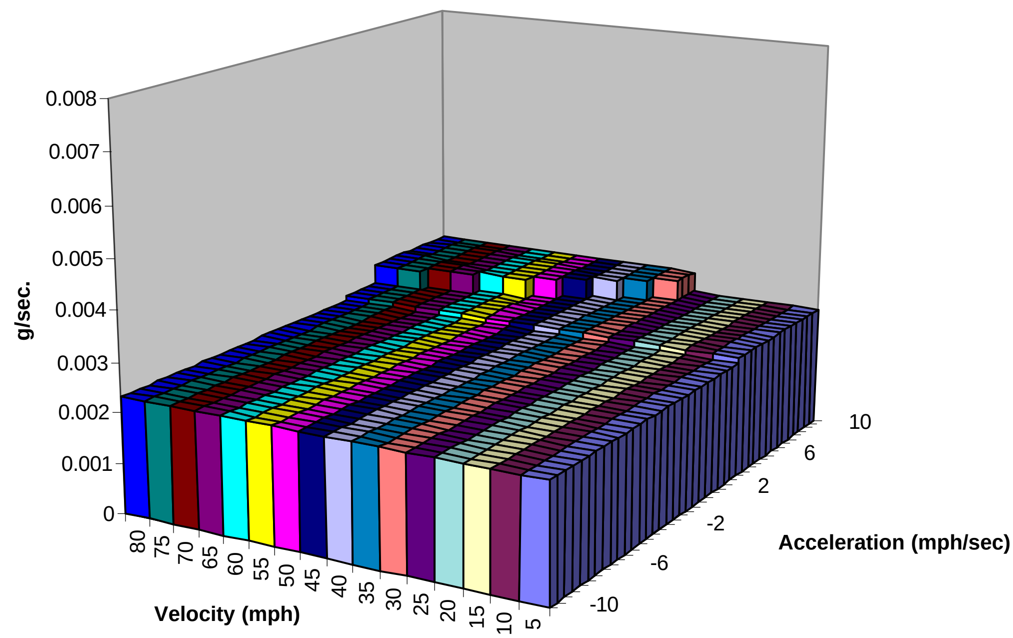



**Figure 6.14 - MEASURE g/sec HC emission rates by velocity and acceleration for the study area's vehicle fleet**

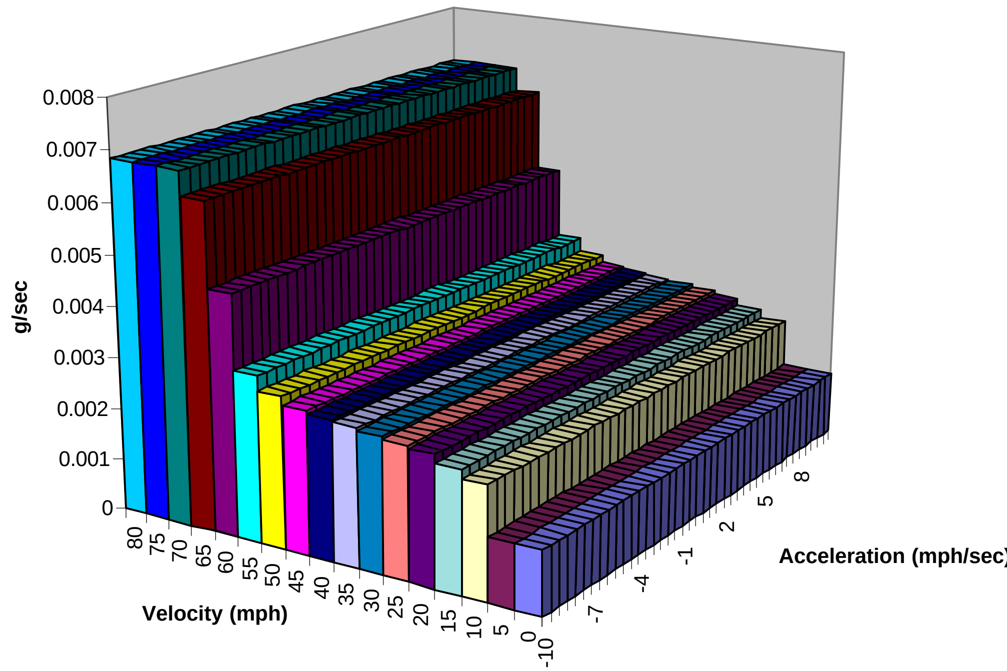



**Figure 6.15 - MOBILE5a g/sec HC emission rates by velocity and acceleration for the sample area's vehicle fleet**

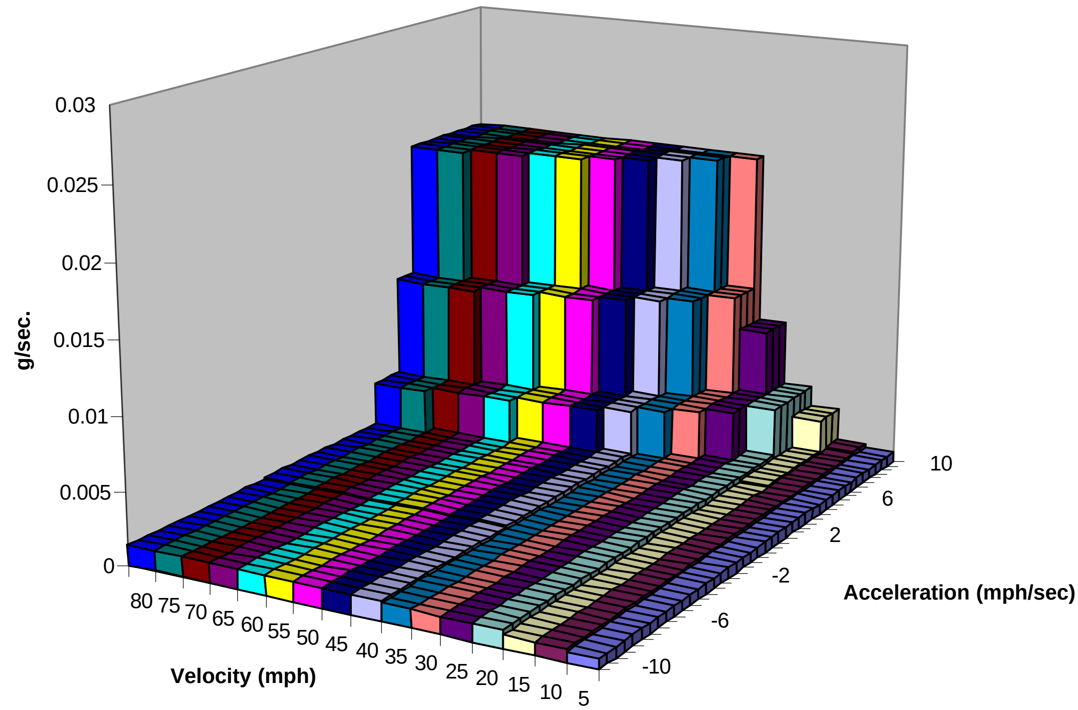



**Figure 6.16 - MEASURE g/sec NOx emission rates by velocity and acceleration for the sample area's vehicle fleet**

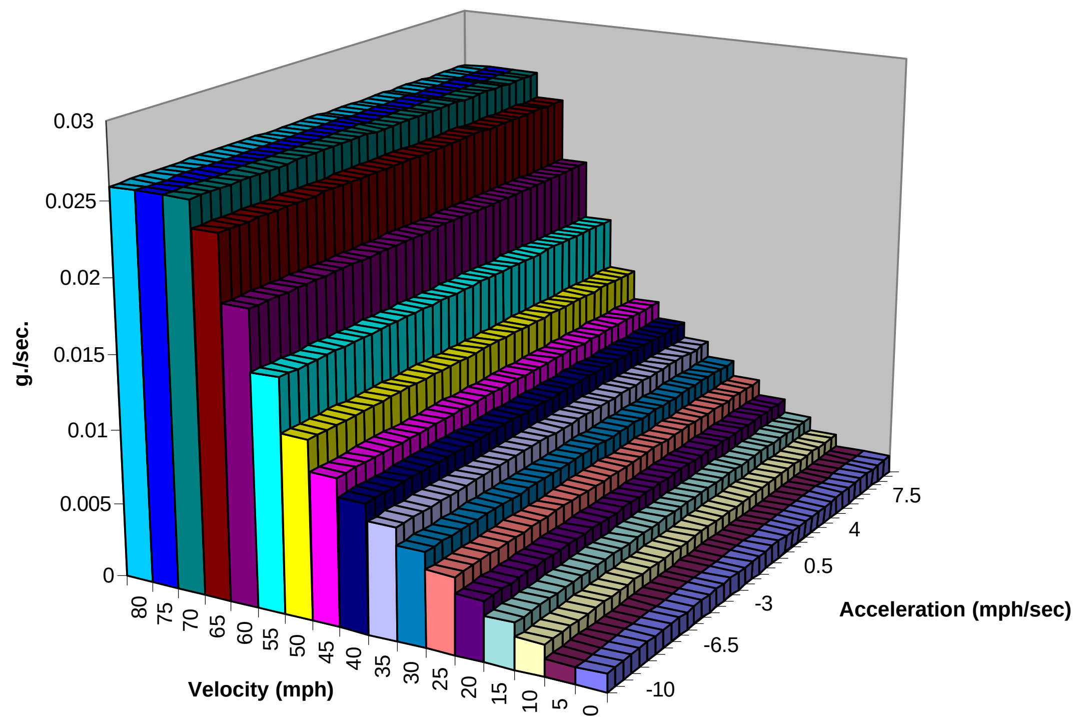



**Figure 6.17 - MOBILE5a g/sec NOx emission rates by velocity and acceleration for the sample area's vehicle fleet**

### 6.4.5. Conclusion

There are a variety of errors generated by model procedures. Future validation efforts will quantify the errors so that confidence bands can be predicted for the estimate value and position. During the process, particular attention should be paid to the estimates of the spatial variability of the operating fleet. Clearly, this model component has the greatest potential for spatial error. The use of regional temporal factors create significant non-spatial errors, particularly in off-peak hours. As long as accurate information is fed to the model, errors resulting from modeling procedures are significantly reduced (polygon overlay error, trip disaggregation error, etc.). Minimum grid cell sizes should be assessed for each model scenario.

The new modal emission rates indicate that vehicle technologies and vehicle operating profiles (speed and acceleration) have significant impacts on emission rates. While the new emission rate models need to be validated, there is strong evidence that MOBILE5a is insensitive to important emission-specific vehicle activity.